ATD

J. gen. Virol. (1983), 64, 997–1006. *Printed in Great Britain*

Key words: *HSV-1/mRNA location/DNA sequence*

Organization of the Herpes Simplex Virus Type 1 Transcription Unit Encoding Two Early Proteins with Molecular Weights of 140000 and 40000

By JOHN McLAUCHLAN AND J. BARKLIE CLEMENTS*

Institute of Virology, University of Glasgow, Glasgow G11 5JR, U.K.

(*Accepted 21 December 1982*)

SUMMARY

The 5′ ends of two early herpes simplex virus type 1 mRNAs have been identified by nuclease S1 and exonuclease VII analysis using cloned virus DNA probes. These mRNAs (5·0 kb and 1·2 kb), located within the genome region between map coordinates 0·56 and 0·60, are unspliced and share a 3′ terminus. Genomic DNA at the 5′ ends has been sequenced and the 5′ termini have been located on the virus DNA sequence. The DNA sequence has revealed signals involved in the initiation of transcription of both mRNAs, and the 5′ end of the 1·2 kb mRNA is encoded within the internal sequences of the 5·0 kb mRNA. The probable translational initiation codons for the polypeptides specified by these mRNAs have been identified, and the results indicate that the coding regions of the two mRNAs do not overlap.

INTRODUCTION

Studies on the transcription of adenoviruses and papovaviruses have revealed that many of their transcription units generate mRNA families which are both 5′ and 3′ co-terminal. Within such transcription units, the various different mRNAs are generated by differential splicing of the primary transcripts (Ziff, 1981).

Transcription of herpes simplex virus (HSV) has been less extensively studied; however, the majority of mRNAs so far examined are unspliced (Anderson *et al.*, 1981; Costa *et al.*, 1981; McKnight, 1980; F. J. Rixon & J. B. Clements, unpublished results). The exceptions are two HSV-1 immediate-early mRNAs which have a common 5′ portion containing an intron located 5′ to their translational initiation codons (Rixon & Clements, 1982; Watson *et al.*, 1981), and the mRNAs from a single, late HSV-1 transcription unit which have also been reported to be spliced (Frink *et al.*, 1981*a*).

Previously, we have analysed the structures of two overlapping HSV-1 mRNAs [5·0 kb (kilobases) and 1·2 kb] which map in the *Hin*dIII *k*/*Bam*HI *o* region of the genome (McLauchlan & Clements, 1982). These mRNAs share a 3′ terminus and have 3′ unspliced portions which extend 770 bases to the right of the *Hin*dIII cleavage site at map coordinate 0·586 (Fig. 1). Both mRNAs are synthesized at very early times post-infection and they exhibit a similar temporal pattern of appearance and disappearance in the cytoplasm.

The genome region at which these HSV-1 mRNAs map is of interest since it corresponds to an HSV-2 region, containing the *Bgl*II *n* fragment, which has been reported to cause morphological transformation of cells *in vitro* (Reyes *et al.*, 1979). While the *Hin*dIII *k*/*Bam*HI *o* region of the HSV-1 genome appears not to transform cells *in vitro* (Reyes *et al.*, 1979), the polypeptides specified by the equivalent HSV-1 and HSV-2 genome regions are similar in size (Anderson *et al.*, 1981; Galloway *et al.*, 1982).

Here, we locate the 5′ termini of the overlapping 5·0 kb and 1·2 kb mRNAs by nuclease S1 and exonuclease VII digestions using 5′-labelled DNA probes. Genomic DNA around the locations of the 5′ ends has been sequenced, and signals involved in the initiation of transcription have been identified. Putative polypeptide coding regions of the mRNAs have also been identified.

0022-1317/83/0000-5444 $02.00

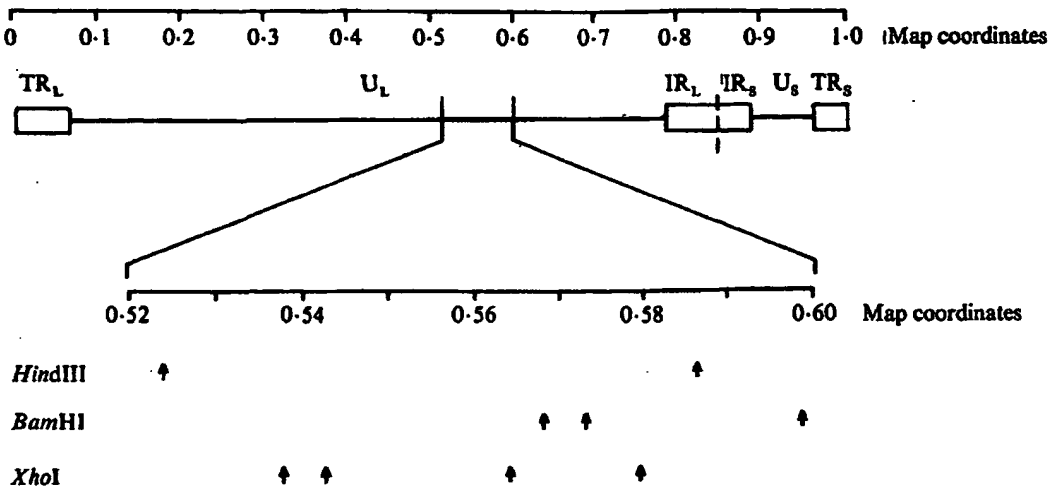

Fig. 1. Restriction endonuclease cleavage maps of HSV-1 strain 17 DNA at the genome region located between map coordinates 0·52 and 0·60.

METHODS

Cells and virus. Baby hamster kidney 21 (C13) cells were grown as monolayers in 800 ml plastic tissue culture flasks (Clements *et al.*, 1977). For the production of early RNA (3 h post-infection), cell monolayers were infected with HSV-1 (Glasgow strain 17) at 10 p.f.u./cell. This was increased to 50 p.f.u./cell to produce immediate-early RNA and cycloheximide-released RNA. As appropriate, cell monolayers were pretreated and maintained in medium containing cycloheximide as described previously (Clements *et al.*, 1977). For isolation of cycloheximide-released RNA, the cycloheximide was removed by washing the cells three times with cycloheximide-free medium at 37 °C. Infection was then continued for 1 h in cycloheximide-free medium at 37 °C, after which the RNA was isolated.

Cell fractionation and isolation of RNA. Cytoplasmic cell fractions were prepared and RNA was isolated as described previously (Kumar & Lindberg, 1972).

Enzymes. All enzymes were obtained from Bethesda Research Laboratories or New England Biolabs, with the exception of T4 polynucleotide kinase (P-L Biochemicals) and nuclease S1 (Boehringer). DNA was digested with restriction endonucleases at 37 °C in 50 to 200 µl 6 mM-Tris–HCl pH 7·5, 6 mM-$MgCl_2$ and 6 mM-2-mercaptoethanol.

Cloning procedures. Fragments of HSV-1 DNA, generated using restriction endonucleases, were cloned within the Institute of Virology under Category I containment conditions (U.K. Genetic Manipulation Advisory Group). The host bacterium was *Escherichia coli* K12 HB101 and the cloning vector was pAT153 (Twigg & Sherratt, 1980). Isolation of cloned virus DNA was as described by Davison & Wilkie (1981).

Purification and end-labelling of DNA fragments. Purification of DNA fragments from agarose or polyacrylamide gels, and labelling of the 5′ and 3′ ends was carried out as described by McLauchlan & Clements (1982).

In order to generate fragments with uniquely labelled ends, the DNA fragments, either 5′- or 3′-labelled at both ends, were redigested with a second restriction endonuclease.

Structural analysis of mRNAs. Structural analysis of mRNAs was performed using the nuclease S1 and exonuclease VII digestion procedures of Berk & Sharp (1978), modified by using either 5′- or 3′-end-labelled DNA probes instead of uniformly labelled DNA (Weaver & Weissmann, 1979).

Either 5- or 3′-labelled DNA (less than 1 µg) was co-precipitated with 50 µg of cytoplasmic RNA from infected or mock-infected cells. The DNA/RNA pellet was resuspended in 20 µl of 90% (v/v) formamide (deionized with Amberlite monobed resin MB-1), 400 mM-NaCl, 40 mM-PIPES pH 6·8, 1 mM-EDTA. This mixture was heated to 90 °C for 3 min then incubated at 57 °C or 57·5 °C for either 5 h or 16 h. Prior to nuclease treatment, the hybridization mixtures were rapidly quenched on ice.

Nuclease S1 digestion was performed at 37 °C for 1 h in 200 µl of 250 mM-NaCl, 30 mM-sodium acetate pH 4·5, 1 mM-$ZnSO_4$ with 4000 units of nuclease S1. The nuclease S1-digested hybrids were extracted with phenol/chloroform then precipitated with ethanol. The digestion products were analysed by gel electrophoresis.

Exonuclease VII digestion was performed at 37 °C for 1 h in 200 µl of 6·7 mM-potassium phosphate pH 7·9, 8·3 mM-EDTA, 10 mM-2-mercaptoethanol with 0·5 units of exonuclease VII. The exonuclease VII-digested hybrids were extracted with phenol/chloroform, desalted on a Pharmacia PD-10 Sephadex G25M column and precipitated with ethanol. The digestion products were analysed by gel electrophoresis.

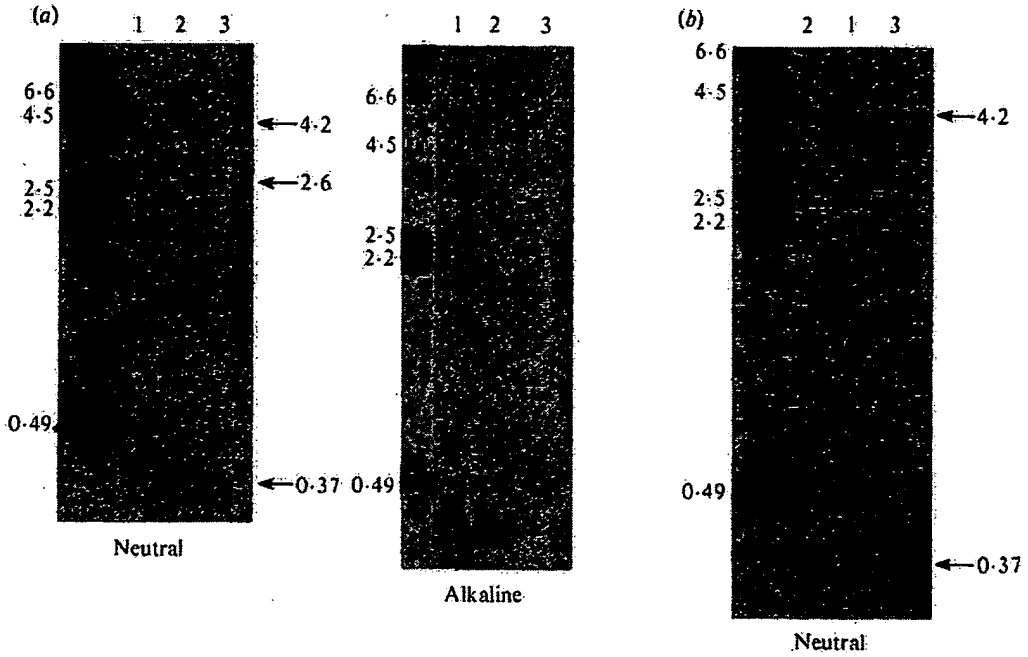

Fig. 2. Analysis of the 5′ ends of mRNAs located in *Hin*dIII *k*. (*a*) The DNA probe was *Hin*dIII *k*, 5′-labelled at both *Hin*dIII sites (Fig. 1, coordinates 0·524 and 0·586). (*b*) The DNA probe was *Hin*dIII *k* uniquely 5′-labelled at one *Hin*dIII site (Fig. 1, coordinate 0·586). The RNA samples used in (*a*) and (*b*) were as follows: 1, immediate-early cytoplasmic RNA; 2, cycloheximide-released cytoplasmic RNA; 3, mock-infected cytoplasmic RNA. The nuclease S1-resistant material was subjected to electrophoresis together with lambda DNA/*Hin*dIII fragment markers on 1·5% (w/v) neutral agarose gels and on a 1·5% (w/v) alkaline agarose gel.

Due to the processive nature of exonuclease VII activity (Chase & Richardson, 1974), this enzyme will leave several undigested nucleotides at a hybrid end. This accounts for the slightly larger size of the exonuclease VII-resistant bands as compared to the equivalent nuclease S1-resistant band.

Gel electrophoresis. Samples were electrophoresed either on non-denaturing 1·5% (w/v) agarose gels in a buffer containing 90 mM-Tris, 90 mM-boric acid pH 8·3, 1 mM-EDTA or on alkaline 1·5% (w/v) agarose gels in 30 mM-NaOH, 2 mM-EDTA. Electrophoresis was carried out at room temperature for 16 h at 50 V. All gels were then dried down and the bands visualized by autoradiography at −70 °C using Kodak X-Omat-S film.

Denaturing polyacrylamide gels, essentially as described by Maxam & Gilbert (1980), were run in 90 mM-Tris, 90 mM-boric acid pH 8·3, 1 mM-EDTA and the gels contained 9 M-urea. Samples were dissolved in deionized formamide and denatured at 90 °C for 2 min before loading. Electrophoresis was carried at room temperature for 3 to 6 h at 40 W. The radiolabelled bands were detected by autoradiography.

DNA sequencing. DNA sequences were determined by chemical degradation (Maxam & Gilbert, 1980) of 5′- and 3′-labelled DNA fragments.

RESULTS

5′ termini of mRNAs mapping in HindIII k

The 5′ portions of the mRNAs were investigated using *Hin*dIII *k* (coordinates 0·524 to 0·586) which was 5′-labelled at both ends. This DNA probe was hybridized to infected and mock-infected RNAs and, after nuclease S1 digestion, the products were separated on 1·5% (w/v) neutral and alkaline agarose gels (Fig. 2*a*).

Three protected DNA fragments of 4·2, 2·6 and 0·37 kb were detected with the cycloheximide-released infected cell RNA (Fig. 2*a*, lane 2) and the 0·37 kb product was the most abundant. These sizes were similar on neutral and alkaline gels, suggesting that those portions of the three mRNAs mapping in *Hin*dIII *k* were unspliced.

The 5′ ends of these mRNAs were orientated with respect to each end of *Hin*dIII *k* by hybridizations with the uniquely labelled (at coordinate 0·586) larger fragment derived by

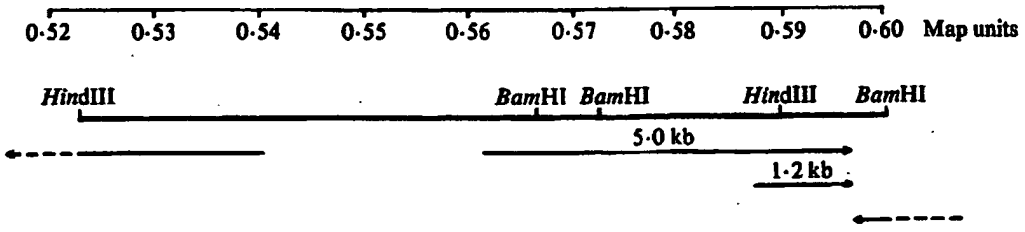

Fig. 3. Summary of the genome map locations and orientations of the HSV-1 mRNAs mapping between coordinates 0·525 and 0·60. The mRNA transcribed leftwards across coordinate 0·60 has been described previously (McLauchlan & Clements, 1982).

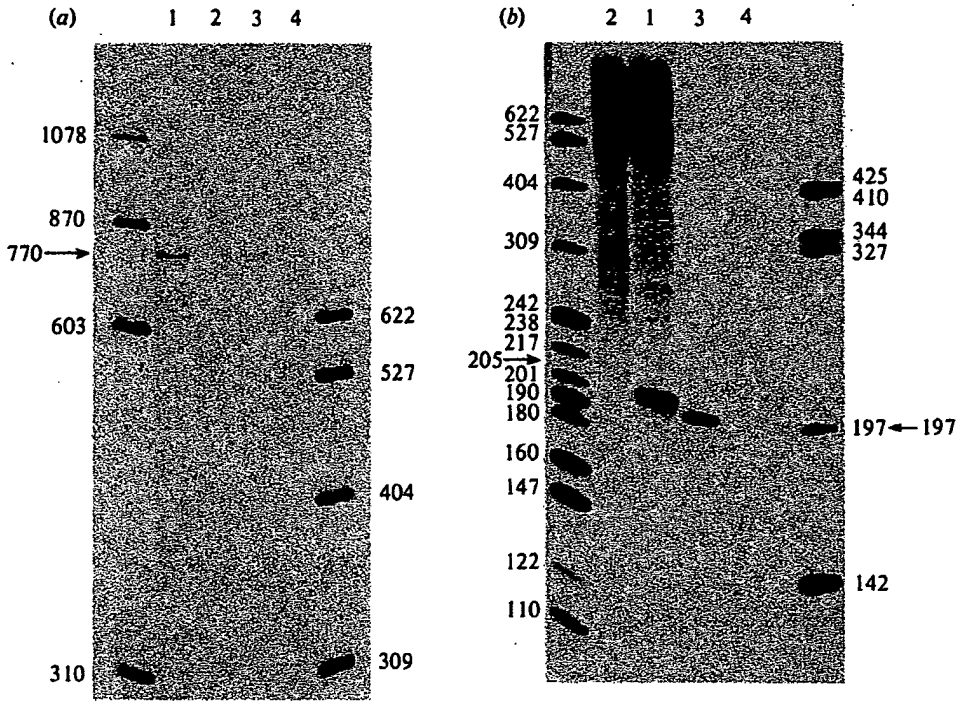

Fig. 4. Precise map location of the 5′ end of the 5·0 kb mRNA. (*a*) The DNA probe was a *Bam*HI/*Hin*dIII fragment (Fig. 1, coordinates 0·524 to 0·568), uniquely 5′-labelled at the *Bam*HI site. (*b*) An *Xho*I fragment probe (Fig. 1, coordinates 0·542 to 0·563) was 5′-labelled at both ends. The RNA samples used in (*a*) and (*b*) were as follows: lanes 1 and 3, 3 h infected cell cytoplasmic RNA; lanes 2 and 4, mock-infected cytoplasmic RNA. Samples 1 and 2 were digested with exonuclease VII and samples 3 and 4 were treated with nuclease S1. The nuclease-resistant material was electrophoresed on 8% denaturing polyacrylamide gels. The size standards used were: ϕX174 DNA/*Hin*cII fragments and pBR322 DNA/*Hpa*II fragments in (*a*); pBR322 DNA/*Hpa*II fragments and 3′-labelled *Hin*fI fragments of pBR322 DNA after digestion with *Hae*III in (*b*).

cleavage of 5′-labelled *Hin*dIII *k* with *Hpa*I (coordinate 0·528). By using this probe, two major nuclease S1-resistant bands of 4·2 kb and 0·37 kb were detected on neutral gels (Fig. 2*b*, lane 2), thus locating the 5′ ends of two mRNAs which are transcribed rightwards across the *Hin*dIII cleavage site at coordinate 0·586 (Fig. 3). The almost complete disappearance of the 2·6 kb band using the *Hpa*I-cleaved *Hin*dIII probe implies that a leftwards-transcribed mRNA has its 5′ end located 2·6 kb from the *Hin*dIII cleavage site at coordinate 0·525 (Fig. 3).

Previously, we have shown that the 3′ ends of the two rightwards-transcribed mRNAs are located 770 bases into *Hin*dIII *l* (coordinates 0·586 to 0·640; McLauchlan & Clements, 1982). Therefore, the total sizes of these mRNAs are approximately 5·0 kb and 1·2 kb.

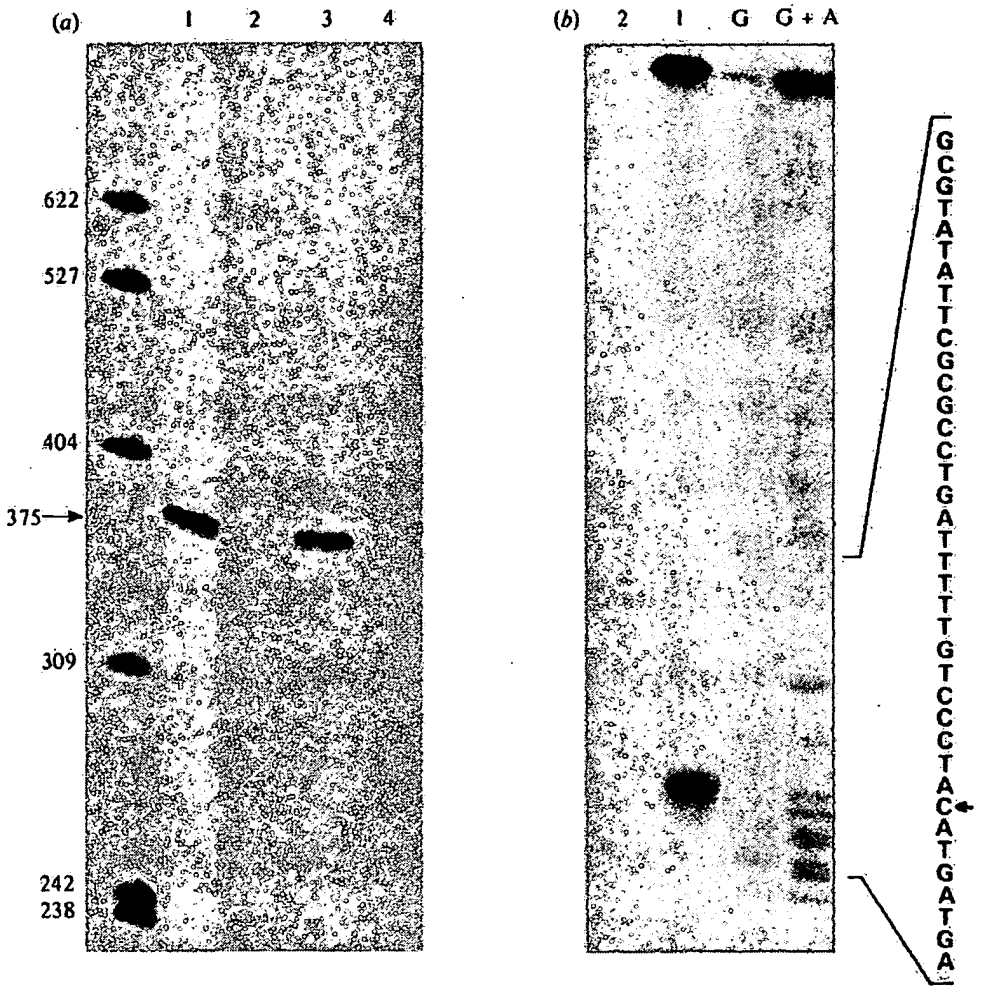

Fig. 5. (*a*) Precise map location of the 5′ end of the 1·2 kb mRNA. An *Hpa*I/*Hin*dIII fragment (coordinates 0·528 to 0·586), uniquely 5′-labelled at the *Hin*dIII site at 0·586, was hybridized to RNA samples: lanes 1 and 3, 3 h infected cell cytoplasmic RNA; lanes 2 and 4, mock-infected cell cytoplasmic RNA. Samples 1 and 2 were digested with nuclease S1 and samples 3 and 4 were treated with exonuclease VII. The nuclease-resistant material was subjected to electrophoresis on an 8% denaturing polyacrylamide gel together with pBR322 DNA/*Hpa*II markers. (*b*) An 8% polyacrylamide sequencing gel which locates the 5′ terminus of the 1·2 kb mRNA on the DNA sequence of the strand coding for the 1·2 kb mRNA. *Bam*HI *o* was uniquely 5′-labelled at a *Hin*fI site (Fig. 8, position 527). This probe was used for sequencing and also was hybridized to RNA samples: lane 1, 3 h infected cell cytoplasmic RNA; lane 2, mock-infected cytoplasmic RNA. Samples 1 and 2 were digested with nuclease S1.

5′ terminus of the 5·0 kb mRNA

The 5′ end of this mRNA was located using a *Hin*dIII/*Bam*HI subclone of *Hin*dIII *k* (Fig. 1, coordinates 0·525 to 0·568). This fragment was uniquely 5′-labelled at the *Bam*HI site then hybridized to infected and mock-infected RNA samples. Following nuclease S1 or exonuclease VII treatment, samples were electrophoresed on an 8% denaturing polyacrylamide gel (Fig. 4*a*).

The single protected band of 770 nucleotides was detected with the infected cell RNA only. The size was similar in the exonuclease VII- and nuclease S1-treated samples (Fig. 4*a*, lanes 1 and 3 respectively), indicating that the 5′ portion of the 5·0 kb mRNA was unspliced.

A more precise location of the 5′ terminus was obtained using a 5′-labelled *Xho*I fragment (Fig. 1, coordinates 0·542 to 0·564). A nuclease S1-resistant product of 197 nucleotides was observed with infected cell RNA; however, the exonuclease VII-resistant product was about 205 nucleotides (Fig. 4*b*, lanes 1 and 3). This small size difference is due to the processive nature of

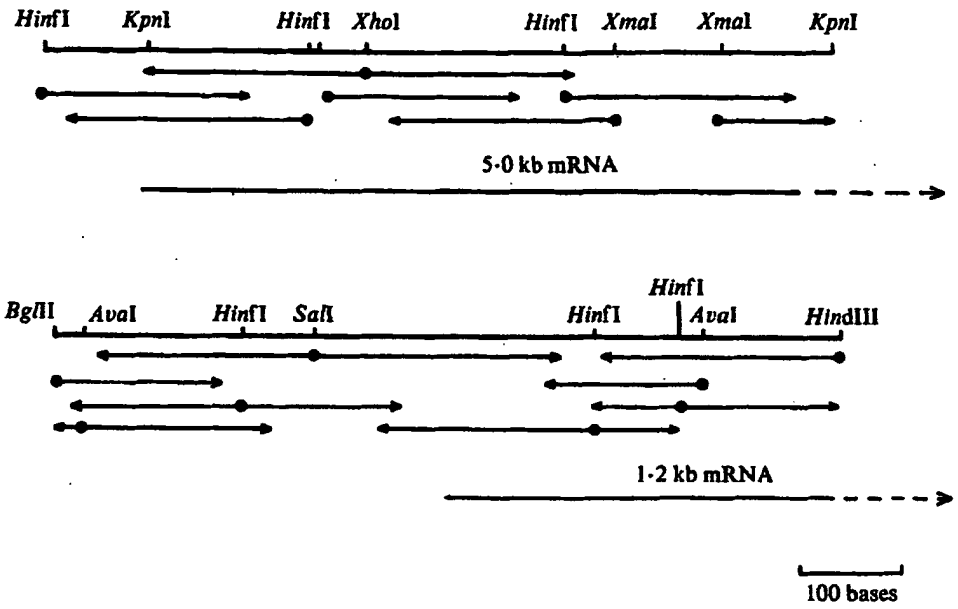

Fig. 6. The restriction endonuclease sites which were used to determine the nucleotide sequences shown in Fig. 7 (5·0 kb mRNA) and Fig. 8 (1·2 kb mRNA).

exonuclease VII activity as explained in Methods. Therefore the 5′ terminus is located 197 nucleotides from a *Xho*I site (Fig. 7, position 289), as determined by the nuclease S1-resistant product.

5′ terminus of the 1·2 kb mRNA

The 5′ end was located using 5′-labelled *Hin*dIII *k*. This fragment was hybridized with RNA extracted from cells 3 h after infection, and with mock-infected cell RNA. A band of 375 nucleotides was detected in both nuclease S1- and exonuclease VII-digested RNA from infected cells (Fig. 5*a*, lanes 1 and 3). This located the 5′ end at 375 nucleotides from the *Hin*dIII site (Fig. 1, coordinate 0·586) and indicated that the 5′ portion of the 1·2 kb mRNA was unspliced.

The 5′ end was more precisely located within a *Hinf*I/*Sal*I fragment (Fig. 8, positions 252 to 527) of *Bam*HI *o*, using a fragment which was uniquely 5′-labelled at the *Hinf*I site. Following hybridization and nuclease S1 treatment, the samples were electrophoresed on an 8% polyacrylamide sequencing gel along with the G and G+A sequence reaction products of the DNA probe (Fig. 5*b*). The sequence on the gel is that of the DNA strand complementary to the 5·0 kb and 1·2 kb mRNAs. The 5′ end extended 157 bases from the *Hinf*I site and the precise location of the 5′ terminus is indicated on the DNA sequence (Fig. 8).

Nucleotide sequences at the 5′ termini of the 5·0 kb and 1·2 kb mRNAs

Sequences were determined by the chemical method (Maxam & Gilbert, 1980) using cloned virus DNA fragments which were uniquely 5′- or 3′-labelled at restriction endonuclease cleavage sites. The restriction endonuclease sites which were used to determine the nucleotide sequences are shown in Fig. 6. All of the sequences shown were obtained for both strands of the DNA, and the sequence data shown in Fig. 8 were determined using DNA isolated from two independently derived clones.

The 5′ end of the 5·0 kb mRNA, located at position 93 (Fig. 7), is positioned at the first adenosine residue within the sequence GTACCA. Similarly, the 5′ end of the 1·2 kb mRNA is located at the guanosine residue within the sequence ATGTAC (Fig. 8, position 371). Sequences at both these termini resemble the 'cap' site sequence found for other eukaryotic genes from diverse sources (Busslinger *et al.*, 1980). 'TATA' box sequences (Corden *et al.*, 1980) are located at positions −30 to −23 (5·0 kb mRNA) and at positions −28 to −22 (1·2 kb mRNA). At

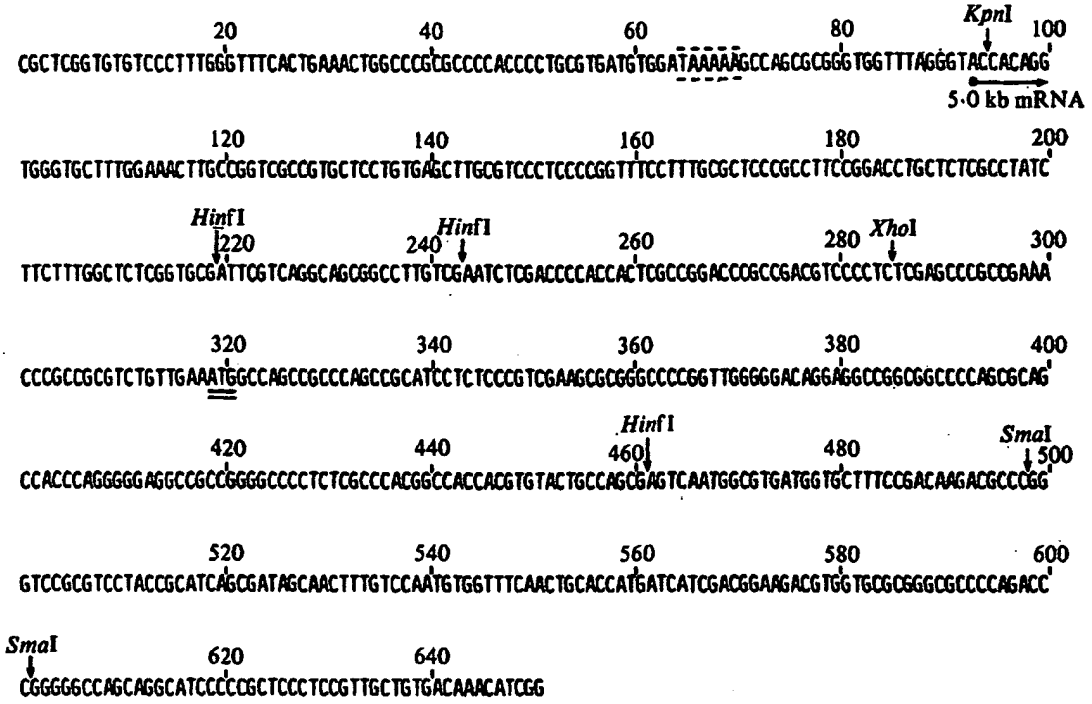

Fig. 7. Nucleotide sequence at the 5′ end of the 5·0 kb mRNA. 'TATA' box sequences are indicated by dotted lines. The location of the 5′ terminus is shown and the first ATG triplet is underlined.

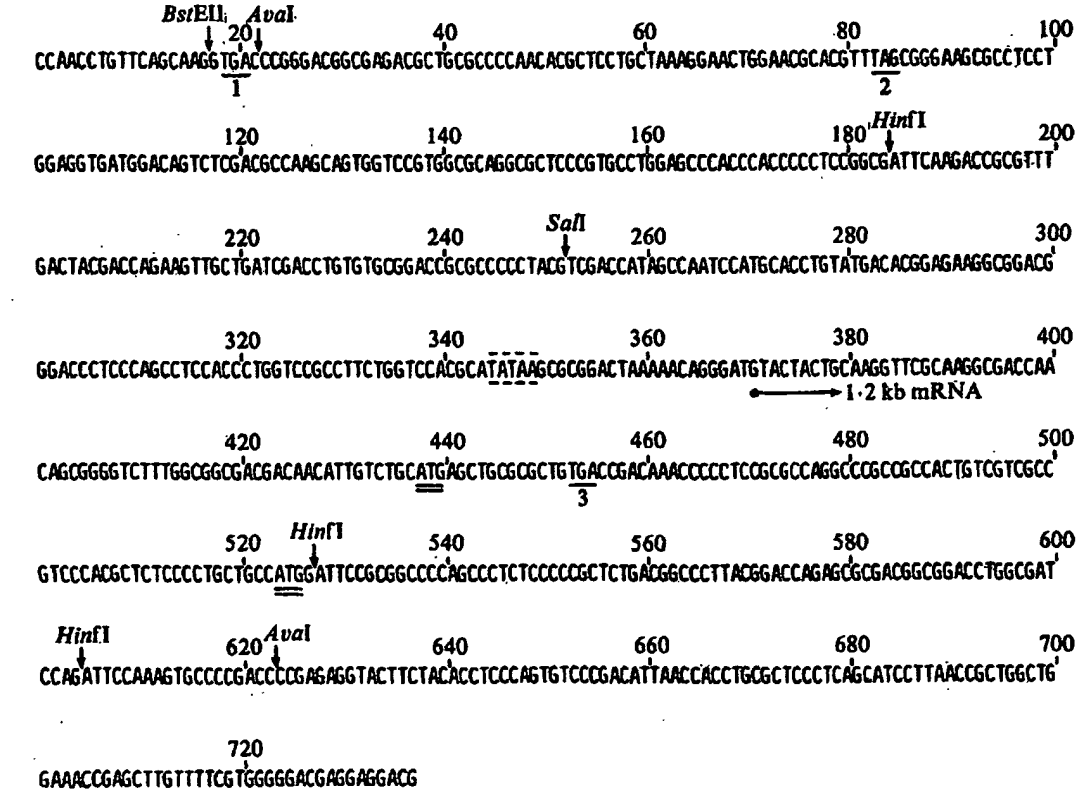

Fig. 8. Nucleotide sequence at the 5′ end of the 1·2 kb mRNA. 'TATA' box sequences are indicated by dotted lines and the 5′ terminus is shown. The first and second ATG triplets encoded by the 1·2 kb mRNA are indicated by two solid lines. The first stop codon in each reading frame within the sequence is underlined and numbered.

32 bp (base pairs) and also at 46 bp upstream from the 5′ end of the 1·2 kb mRNA is the pentameric sequence GGTCC which closely resembles the consensus sequence GATCC, sometimes observed at eukaryotic promoter regions (Busslinger *et al.*, 1980); no close homology to this sequence is found upstream from the 5′ end of the 5·0 kb mRNA.

Coding regions of the 5·0 kb and 1·2 kb mRNAs

The 5·0 kb mRNA, by *in vitro* translation, specified a polypeptide of 140000 mol. wt. (Anderson *et al.*, 1981). The first ATG triplet is located 227 bases downstream from the 5′ end of the 5·0 kb mRNA (Fig. 7, position 319) and this initiation codon lies in the only open reading frame which extends for 332 bases, the limit of our data. The C-terminal location of the 140000 polypeptide is unknown; however, an open reading frame of 452 bases in our sequence data extends to 82 bases downstream from the 5′ end of the 1·2 kb mRNA (Fig. 8, position 453).

The 1·2 kb mRNA, by *in vitro* translation, appeared to specify a polypeptide of 40000 mol. wt. (Anderson *et al.*, 1981). The first ATG triplet within the sequences of the 1·2 kb mRNA is located 68 bases downstream from the 5′ end (Fig. 8, position 438); however, this reading frame is closed after 5 codons. A second ATG triplet is positioned 86 bases further downstream (Fig. 8, position 524) from the first ATG triplet, and this codon lies in an open reading frame which extends throughout the next 214 nucleotides to the limit of our sequence.

DISCUSSION

We have analysed the 5′ portions of two overlapping mRNAs (5·0 kb and 1·2 kb), specified by HSV-1 strain 17, which map at the *Hin*dIII *k/l* region of the genome and have located their 5′ termini on the genomic DNA sequence. Nuclease S1 and exonuclease VII analyses indicate that the 5′ portions are unspliced. The mRNAs share a 3′ terminus located in *Hin*dIII *l*, and their 3′ portions also are unspliced.

Anderson *et al.* (1981) have described the 5·0 kb and 1·2 kb mRNAs in cells infected with HSV-1 strain KOS and also have mapped 7·0 kb and 1·5 kb mRNAs within this region. These additional mRNAs appeared to have common 5′ ends and they suggested that the 7·0 kb mRNA was 3′ co-terminal with the 5·0 kb and 1·2 kb mRNAs.

We did not detect the 7·0 kb and 1·5 kb mRNAs; however, this may be a reflection of virus or cell strain differences. The relative abundance of individual mRNAs can vary, as evidenced by comparing the levels of the 5·0 kb and 1·2 kb mRNAs in cells infected with either strain 17 or strain KOS. Strain 17 produces much more of the 1·2 kb mRNA relative to the 5·0 kb species (Fig. 2*b*, lane 2), whereas the 1·2 kb mRNA was only just detectable in cells infected with KOS (Anderson *et al.*, 1981).

The 5·0 kb and 1·2 kb mRNAs comprise a transcription unit which consists of unspliced, 3′ co-terminal mRNAs with different 5′ ends. Signals involved in transcription initiation of the 1·2 kb mRNA (such as the 'TATA' box, cap site and pentameric sequence) are present within the 5·0 kb mRNA sequences. A transcription unit with this type of organization does not fit into either of the two categories outlined by Darnell (1982); it is neither 'simple' (more than one polypeptide is encoded) nor 'complex' (there is no splicing or more than one poly(A) site).

Transcription units with similar arrangements have been described in adenovirus types 2 and 5, and in yeast. The yeast invertase locus specifies two apparently unspliced mRNAs which code for different forms of invertase (Carlson & Botstein, 1982). These mRNAs are 3′ co-terminal and the 5′ end of the 1·8 kb mRNA is located within the sequences specifying the 1·9 kb. In adenovirus, the 'TATA' box for polypeptide IX mRNA, an unspliced message, lies within the intron sequences of the EIb transcription unit (Alestrom *et al.*, 1980); the 5′ end of IVa2 mRNA is located within DNA sequences encoding the EIIb mRNAs (Stillman *et al.*, 1981) and the 'TATA' box for the EIIa late mRNAs is located within a region that is expressed at early times (Chow *et al.*, 1979).

In the adenovirus cases, the 5′ end of a late mRNA is located within a region that is expressed at early times. In contrast, the 5·0 kb and 1·2 kb HSV-1 mRNAs are both early species which appear in the cytoplasm simultaneously (McLauchlan & Clements, 1982) and there is no constraint on initiation of the 1·2 kb mRNA.

The nucleotide sequences at the 5′ end of the 5·0 kb mRNA presented here largely agree with and extend those of Frink *et al.* (1981*b*). A striking difference is the absence in strain 17 of an apparently tandemly reiterated sequence (CCGCCGAAAC), located 40 bases upstream from the first ATG triplet in strain KOS. Both sets of data locate the 'TATA' box and cap site of the 5·0 kb mRNA in similar positions. However, we place the 5′ terminus one nucleotide further downstream from the 'TATA' box.

The 5·0 kb mRNA encodes a 140000 mol. wt. polypeptide (Anderson *et al.*, 1981), the amino-terminus of which has not been located. The first ATG triplet lies 227 nucleotides downstream from the 5′ end, and is followed by an open reading frame which extends to the limit of our sequence. The 1·2 kb mRNA appears to encode a 40000 mol. wt. polypeptide. However, the reading frame following the first ATG triplet, 68 nucleotides downstream from the 5′ end, is closed after 5 codons. A second ATG triplet, located 154 bases downstream from the 5′ end, lies in an open reading frame that extends as far as we have sequenced.

The sequences flanking both the second ATG of the 1·2 kb mRNA and the first ATG of the 5·0 kb mRNA agree closely with the preferred signals for initiation of translation as described by Kozak (1981). By contrast, the nucleotides flanking the first ATG in the 1·2 kb mRNA resemble those of non-functional initiation codons. The data provide no information on whether either mRNA is functional in translation: it is possible that the 5·0 kb mRNA could specify both polypeptides.

We propose that the amino-terminus of the 40000 mol. wt. polypeptide is specified by the second ATG triplet. Our RNA mapping and sequencing data indicate that the 140000 and 40000 mol. wt. polypeptides do not share coding sequences in common. Thus, the carboxy-terminus of the 140000 mol. wt. polypeptide appears to be located within sequences specifying the 5′ untranslated leader of the 1·2 kb mRNA.

Recently, Draper *et al.* (1982) have published the complete nucleotide sequence of the 1·2 kb mRNA of strain KOS. Comparison of the strain KOS sequence with that of strain 17 presented here reveals a number of differences. Both sets of data indicate that the coding regions of the 140000 mol. wt. and 40000 mol. wt. polypeptides do not overlap; however, all three reading frames of strain KOS apparently terminate upstream of the 5′ end of the 1·2 kb mRNA, in contrast to the situation already described for strain 17. Also, the first ATG in the 1·2 kb mRNA of strain 17 (Fig. 8, position 438) was not found in strain KOS. Finally, there are additional bases between positions 545 and 562 (Fig. 8) in strain 17 which are not present in the reported nucleotide sequence of strain KOS. The presence of these additional bases results in a different amino acid sequence for the 40000 mol. wt. polypeptide between positions 545 and 562 (Fig. 8). Outside this region, the same reading frame is used in both sets of nucleotide sequence. In order to ensure that these differences were not due to sequencing errors in strain 17, the nucleotide sequence was derived from both strands of the DNA using 5′- and 3′-labelled DNA fragments (Fig. 6) and from two independently derived clones.

Furthermore, we have sequenced the equivalent region of HSV-2 (J. McLauchlan & J. B. Clements, unpublished results). The polypeptides in this region have molecular weights of 140000 and 35000 (Galloway *et al.*, 1982). Our results indicate that, as in HSV-1 strain 17, the coding regions for the HSV-2 polypeptides do not overlap and that there is a reading frame which terminates within the sequences encoding the 5′ end of a 1·2 kb mRNA. The nucleotide sequence in this region is highly homologous with that derived from strain 17 and the additional bases between positions 545 and 562 (Fig. 8) in strain 17 also are present in HSV-2.

We would like to thank Frazer J. Rixon for helpful discussion and Professor J. H. Subak-Sharpe for critical reading of the manuscript. This project was supported by MRC project grant AG978/709/S.

REFERENCES

ALESTROM, P., AKUSJARVI, G., PERRICAUDET, M., MATHEWS, M. B., KLESSIG, D. F. & PETTERSSON, U. (1980). The gene for polypeptide IX of adenovirus type 2 and its unspliced messenger RNA. *Cell* 19, 671–682.

ANDERSON, K. P., FRINK, R. J., DEVI, G. B., GAYLORD, B. H., COSTA, R. H. & WAGNER, E. K. (1981). Detailed characterisation of the mRNA mapping in the Hind III *k* region of the herpes simplex virus type 1 genome. *Journal of Virology* 37, 1011–1027.

BERK, A. J. & SHARP, P. A. (1978). Structure of the adenovirus 2 early mRNAs. *Cell* **14**, 695–711.

BUSSLINGER, M., PORTMANN, R., IRMINGER, J. C. & BIRNSTIEL, M. L. (1980). Ubiquitous and gene-specific regulatory 5′ sequences in a sea urchin histone DNA clone coding for histone protein variants. *Nucleic Acids Research* **8**, 957–977.

CARLSON, M. & BOTSTEIN, D. (1982). Two differentially regulated mRNAs with different 5′ ends encode secreted and intracellular forms of invertase. *Cell* **28**, 145–154.

CHASE, J. W. & RICHARDSON, C. C. (1974). Exonuclease VII of *Escherichia coli*. *Journal of Biological Chemistry* **249**, 4553–4561.

CHOW, L. T., BROKER, T. R. & LEWIS, J. B. (1979). Complex splicing patterns of RNAs from the early regions of adenovirus-2. *Journal of Molecular Biology* **134**, 265–303.

CLEMENTS, J. B., WATSON, R. J. & WILKIE, N. M. (1977). Temporal regulation of herpes simplex virus type 1 transcription: location of transcripts on the viral genome. *Cell* **12**, 275–285.

CORDEN, J., WASYLYK, B., BUCHWALDER, A., SASSONE-CORSI, P., KEDINGER C. & CHAMBON, P. (1980). Promoter sequences of eukaryotic protein-coding genes. *Science* **209**, 1406–1414.

COSTA, R. H., DEVI, B. G., ANDERSON, K. P., GAYLORD, B. H. & WAGNER, E. K. (1981). Characterisation of a major late herpes simplex virus type 1 mRNA. *Journal of Virology* **38**, 483–496.

DARNELL, J. E. (1982). Variety in the level of gene control in eukaryotic cells. *Nature, London* **297**, 365–371.

DAVISON, A. J. & WILKIE, N. M. (1981). Nucleotide sequences of the joint between the L and S segments of herpes simplex virus types 1 and 2. *Journal of General Virology* **55**, 315–331.

DRAPER, K. G., FRINK, R. J. & WAGNER, E. W. (1982). Detailed characterisation of an apparently unspliced β herpes simplex virus type 1 gene mapping in the interior of another. *Journal of Virology* **43**, 1123–1128.

FRINK, R. J., ANDERSON, K. P. & WAGNER, E. K. (1981*a*). Herpes simplex virus type 1 Hind III fragment *l* encodes spliced and complementary mRNA species. *Journal of Virology* **39**, 559–572.

FRINK, R. J., DRAPER, K. G. & WAGNER, E. K. (1981*b*). Uninfected cell polymerase efficiently transcribes early but not late herpes simplex virus type 1 mRNA. *Proceedings of the National Academy of Sciences, U.S.A.* **78**, 6139–6143.

GALLOWAY, D. A., GOLDSTEIN, L. C. & LEWIS, J. B. (1982). Identification of proteins encoded by a fragment of herpes simplex virus type 2 DNA that has transforming activity. *Journal of Virology* **42**, 530–537.

KOZAK, M. (1981). Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Research* **9**, 5233–5252.

KUMAR, A. & LINDBERG, U. (1972). Characterisation of messenger ribonucleoprotein and messenger RNA from KB cells. *Proceedings of the National Academy of Sciences, U.S.A.* **69**, 681–685.

McKNIGHT, S. L. (1980). The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. *Nucleic Acids Research* **8**, 5949–5964.

McLAUCHLAN, J. & CLEMENTS, J. B. (1982). A 3′ co-terminus of two early herpes simplex virus type 1 mRNAs. *Nucleic Acids Research* **10**, 501–512.

MAXAM, A. M. & GILBERT, W. (1980). Sequencing end-labelled DNA with base-specific chemical cleavages. *Methods in Enzymology* **65**, 499–560.

REYES, G. R., LAFEMINA, R., HAYWARD, S. D. & HAYWARD, G. S. (1979). Morphological transformation by DNA fragments of human herpesviruses: evidence for two distinct transforming regions in herpes simplex virus types 1 and 2 and lack of correlation with biochemical transfer of the thymidine kinase gene. *Cold Spring Harbor Symposia on Quantitative Biology* **44**, 629–641.

RIXON, F. J. & CLEMENTS, J. B. (1982). Detailed structural analysis of two spliced HSV-1 immediate-early mRNAs. *Nucleic Acids Research* **10**, 2241–2256.

STILLMAN, B. W., LEWIS, J. B., CHOW, L. T., MATHEWS, M. B. & SMART, J. E. (1981). Identification of the gene and mRNA for the adenovirus terminal protein precursor. *Cell* **23**, 497–508.

TWIGG, A. J. & SHERRATT, D. (1980). Trans-complementable copy-number mutants of plasmid ColE1. *Nature, London* **283**, 216–218.

WATSON, R. J., SULLIVAN, M. & VANDE WOUDE, G. F. (1981). Structures of two spliced herpes simplex type 1 immediate-early mRNAs which map at the junctions of the unique and reiterated regions of the virus DNA S component. *Journal of Virology* **37**, 431–444.

WEAVER, R. F. & WEISSMANN, C. (1979). Mapping of RNA by a modification of the Berk–Sharp procedure: the 5′ termini of 15S β-globin mRNA have identical map coordinates. *Nucleic Acids Research* **7**, 1175–1193.

ZIFF, E. B. (1981). Transcription and RNA processing by the DNA tumour viruses. *Nature, London* **287**, 491–499.

(*Received 28 September 1982*)

The Journal of General Virology

Volume 64, Part 5, May 1983

Printed in Great Britain by Spottiswoode Ballantyne Ltd
Colchester and London

THE JOURNAL OF GENERAL VIROLOGY

PUBLISHED BY

THE SOCIETY FOR GENERAL MICROBIOLOGY

Editorial Office

HARVEST HOUSE, 62 LONDON ROAD, READING, RG1 5AS, BERKSHIRE. *Tel.*: (0734) 861345

Subscription Information

The *Journal of General Virology* is published monthly: one volume (64) of 12 parts will be published in 1983.

	U.K. and Elsewhere	*North America*	*Japan*
1 year (12 parts)	£215.00	U.S.$540.00	£230.00
Per part	£20.00	U.S.$48.00	£21.00

Air freight to North America (U.S.A., Canada and Mexico). Accelerated Surface Post to Japan.

Orders and subscriptions should be sent to *Journal of General Virology*, P.O. Box 32, Commerce Way, Colchester CO2 8HP, U.K., or through an agent.

Back Numbers and Lost Issues

Enquiries for back numbers and claims regarding lost or damaged issues should be addressed to *Journal of General Virology*, P.O. Box 32, Commerce Way, Colchester CO2 8HP, U.K.

Advertisements

Enquiries for advertising space should be sent to the Advertisement Manager, 60 Church Crescent, London N20 0JP.

Postmaster: send address changes in U.S.A. and Canada to Expediters of the Printed Word Ltd, Air & Sea Freight Inc., 527 Madison Avenue, New York, N.Y. 10022, U.S.A.

Second class postage paid at New York, N.Y.